

SEQUENCE RECORD

(1) GENERAL INDICATIONS:

- (i) APPLICANT:
 - (A) NAME: Deutsches Krebsforschungszentrum
 - (B) STREET: Im Neuenheimer Feld 280
 - (C) TOWN: Heidelberg
 - (E) COUNTRY: Germany
 - (F) POSTAL CODE: 69120
- (ii) TITLE OF THE INVENTION: Multivalent Antibody Constructs
- (iii) NUMBER OF SEQUENCES: 17
- (iv) COMPUTER-READABLE VERSION:
 - (A) DATA CARRIER: floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, version #1.30 (EPA)

(2) INDICATIONS AS TO SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1698 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: genome DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) POSITION: 28..1689
- (ix) FEATURE:
 - (A) NAME/KEY: mat_peptide
 - (B) POSITION: 28..1689
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTTCATTA AAGAGGAGAA ATTAACC ATG AAA TAC CTA TTG CCT ACG GCA
Met Lys Tyr Leu Leu Pro Thr Ala
1 5

GCC GCT GGC TTG CTG CTG GCA GCT CAG CCG GCC ATG GCG CAG GTG Ala Ala Gly Leu Leu Leu Ala Ala Gln Pro Ala Met Ala Gln Val 10 15 20	99
CAA CTG CAG CAG TCT GGG GCT GAA CTG GCA AGA CCT GGG GCC TCA GTG Gln Leu Gln Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val 25 30 35 40	147
AAG ATG TCC TGC AAG GCT TCT GGC TAC ACC TTT ACT AGG TAC ACG ATG Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Arg Tyr Thr Met- 45 50 55	195
CAC TGG GTA AAA CAG AGG CCT GGA CAG GGT CTG GAA TGG ATT GGA TAC His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile Gly Tyr 60 65 70	243
ATT AAT CCT AGC CGT GGT TAT ACT AAT TAC AAT CAG AAG TTC AAG GAC Ile Asn Pro Ser Arg Gly Tyr Thr Asn Tyr Asn Gln Lys Phe Lys Asp 75 80 85	291
AAG GCC ACA TTG ACT ACA GAC AAA TCC TCC AGC ACA GCC TAC ATG CAA Lys Ala Thr Leu Thr Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln 90 95 100	339
CTG AGC AGC CTG ACA TCT GAG GAC TCT GCA GTC TAT TAC TGT GCA AGA Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg 105 110 115 120	387
TAT TAT GAT GAT CAT TAC AGC CTT GAC TAC TGG GGC CAA GGC ACC ACT Tyr Tyr Asp Asp His Tyr Ser Leu Asp Tyr Trp Gly Gln Gly Thr Thr 125 130 135	435
CTC ACA GTC TCC TCA GCC AAA ACA ACA CCC AAG CTT GGC GGT GAT ATC Leu Thr Val Ser Ser Ala Lys Thr Thr Pro Lys Leu Gly Gly Asp Ile 140 145 150	483
TTG CTC ACC CAA ACT CCA GCT TCT TTG GCT GTG TCT CTA GGG CAG AGG Leu Leu Thr Gln Thr Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg 155 160 165	531
GCC ACC ATC TCC TGC AAG GCC AGC CAA AGT GTT GAT TAT GAT GGT GAT Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val Asp Tyr Asp Gly Asp 170 175 180	579
AGT TAT TTG AAC TGG TAC CAA CAG ATT CCA GGA CAG CCA CCC AAA CTC Ser Tyr Leu Asn Trp Tyr Gln Gln Ile Pro Gly Gln Pro Pro Lys Leu 185 190 195 200	627
CTC ATC TAT GAT GCA TCC AAT CTA GTT TCT GGG ATC CCA CCC AGG TTT Leu Ile Tyr Asp Ala Ser Asn Leu Val Ser Gly Ile Pro Pro Arg Phe 205 210 215	675
AGT GGC AGT GGG TCT GGG ACA GAC TTC ACC CTC AAC ATC CAT CCT GTG Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Asn Ile His Pro Val 220 225 230	723

GAG AAG GTG GAT GCT GCA ACC TAT CAC TGT CAG CAA AGT ACT GAG GAT Glu Lys Val Asp Ala Ala Thr Tyr His Cys Gln Gln Ser Thr Glu Asp 235 240 245	771
CCG TGG ACG TTC GGT GGA GGC ACC AAG CTG GAA ATC AAA CGG GCT GAT Pro Trp Thr Phe Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Asp 250 255 260	819
GCT GCG GCC GCT GGT GGT GGT TCT GGC GGC GGT GGT AGC GGT GGT Ala Ala Ala Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly 265 270 275 280	867
GGC GGC TCC GGT GGT GGT AGC CAG GTG CAG CTG CAG CAG TCT GGG Gly Gly Ser Gly Gly Ser Gln Val Gln Leu Gln Gln Ser Gly 285 290 295	915
GCT GAG CTG GTG AGG CCT GGG TCC TCA GTG AAG ATT TCC TGC AAG GCT Ala Glu Leu Val Arg Pro Gly Ser Ser Val Lys Ile Ser Cys Lys Ala 300 305 310	963
TCT GGC TAT GCA TTC AGT AGC TAC TGG ATG AAC TGG GTG AAG CAG AGG Ser Gly Tyr Ala Phe Ser Ser Tyr Trp Met Asn Trp Val Lys Gln Arg 315 320 325	1011
CCT GGA CAG GGT CTT GAG TGG ATT GGA CAG ATT TGG CCT GGA GAT GGT Pro Gly Gln Gly Leu Glu Trp Ile Gly Gln Ile Trp Pro Gly Asp Gly 330 335 340	1059
GAT ACT AAC TAC AAT GGA AAG TTC AAG GGT AAA GCC ACT CTG ACT GCA Asp Thr Asn Tyr Asn Gly Lys Phe Lys Gly Lys Ala Thr Leu Thr Ala 345 350 355 360	1107
GAC GAA TCC TCC AGC ACA GCC TAC ATG CAA CTC AGC AGC CTA GCA TCT Asp Glu Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Ser Leu Ala Ser 365 370 375	1155
GAG GAC TCT GCG GTC TAT TTC TGT GCA AGA CGG GAG ACT ACG ACG GTA Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg Arg Glu Thr Thr Val 380 385 390	1203
GGC CGT TAT TAC TAT GCT ATG GAC TAC TGG GGT CAA GGA ACC TCA GTC Gly Arg Tyr Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr Ser Val 395 400 405	1251
ACC GTC TCC TCA GCC AAA ACA ACA CCC AAG CTT GGC GGT GAT ATC GTG Thr Val Ser Ser Ala Lys Thr Thr Pro Lys Leu Gly Gly Asp Ile Val 410 415 420	1299
CTC ACT CAG TCT CCA GCA ATC ATG TCT GCA TCT CCA GGG GAG AAG GTC Leu Thr Gln Ser Pro Ala Ile Met Ser Ala Ser Pro Gly Glu Lys Val 425 430 435 440	1347
ACC ATG ACC TGC AGT GCC AGC TCA AGT GTA AGT TAC ATG AAC TGG TAC Thr Met Thr Cys Ser Ala Ser Ser Val Ser Tyr Met Asn Trp Tyr 445 450 455	1395

CAG CAG AAG TCA GGC ACC TCC CCC AAA AGA TGG ATT TAT GAC ACA TCC Gln Gln Lys Ser Gly Thr Ser Pro Lys Arg Trp Ile Tyr Asp Thr Ser 460 465 470	1443
AAA CTG GCT TCT GGA GTC CCT GCT CAC TTC AGG GGC AGT GGG TCT GGG Lys Leu Ala Ser Gly Val Pro Ala His Phe Arg Gly Ser Gly Ser Gly 475 480 485	1491
ACC TCT TAC TCT CTC ACA ATC AGC GGC ATG GAG GCT GAA GAT GCT GCC Thr Ser Tyr Ser Leu Thr Ile Ser Gly Met Glu Ala Glu Asp Ala Ala 490 495 500	1539
ACT TAT TAC TGC CAG CAG TGG AGT AGT AAC CCA TTC ACG TTC GGC TCG Thr Tyr Tyr Cys Gln Gln Trp Ser Ser Asn Pro Phe Thr Phe Gly Ser 505 510 515 520	1587
GGG ACA AAG TTG GAA ATA AAC CGG GCT GAT ACT GCA CCA ACT GGA TCC Gly Thr Lys Leu Glu Ile Asn Arg Ala Asp Thr Ala Pro Thr Gly Ser 525 530 535	1635
GAA CAA AAG CTG ATC TCA GAA GAA GAC CTA AAC TCA CAT CAC CAT CAC Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ser His His His His 540 545 550	1683
CAT CAC TAATCTAGA His His	1698

(2) INDICATIONS AS TO ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 554 amino acids
 - (B) KIND: amino acid
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
 1 5 10 15

Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Gln Ser Gly Ala Glu
 20 25 30

Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly
 35 40 45

Tyr Thr Phe Thr Arg Tyr Thr Met His Trp Val Lys Gln Arg Pro Gly
 50 55 60

Gln Gly Leu Glu Trp Ile Gly Tyr Ile Asn Pro Ser Arg Gly Tyr Thr
 65 70 75 80

Asn Tyr Asn Gln Lys Phe Lys Asp Lys Ala Thr Leu Thr Thr Asp Lys
 85 90 95

Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp
 100 105 110

Ser Ala Val Tyr Tyr Cys Ala Arg Tyr Tyr Asp Asp His Tyr Ser Leu
 115 120 125

Asp Tyr Trp Gly Gln Gly Thr Thr Leu Thr Val Ser Ser Ala Lys Thr
 130 135 140

Thr Pro Lys Leu Gly Gly Asp Ile Leu Leu Thr Gln Thr Pro Ala Ser
 145 150 155 160

Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser
 165 170 175

Gln Ser Val Asp Tyr Asp Gly Asp Ser Tyr Leu Asn Trp Tyr Gln Gln
 180 185 190

Ile Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr Asp Ala Ser Asn Leu
 195 200 205

Val Ser Gly Ile Pro Pro Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
 210 215 220

Phe Thr Leu Asn Ile His Pro Val Glu Lys Val Asp Ala Ala Thr Tyr
 225 230 235 240

His Cys Gln Gln Ser Thr Glu Asp Pro Trp Thr Phe Gly Gly Thr
 245 250 255

Lys Leu Glu Ile Lys Arg Ala Asp Ala Ala Ala Ala Gly Gly Gly
 260 265 270

Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
 275 280 285

Gln Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ser
 290 295 300

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ala Phe Ser Ser Tyr
 305 310 315 320

Trp Met Asn Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile
 325 330 335

Gly Gln Ile Trp Pro Gly Asp Gly Asp Thr Asn Tyr Asn Gly Lys Phe
 340 345 350

Lys Gly Lys Ala Thr Leu Thr Ala Asp Glu Ser Ser Ser Thr Ala Tyr
 355 360 365

Met Gln Leu Ser Ser Leu Ala Ser Glu Asp Ser Ala Val Tyr Phe Cys
 370 375 380
 Ala Arg Arg Glu Thr Thr Val Gly Arg Tyr Tyr Tyr Ala Met Asp
 385 390 395 400
 Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Lys Thr Thr
 405 410 415
 Pro Lys Leu Gly Gly Asp Ile Val Leu Thr Gln Ser Pro Ala Ile Met
 420 425 430
 Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Ser Ala Ser Ser
 435 440 445
 Ser Val Ser Tyr Met Asn Trp Tyr Gln Gln Lys Ser Gly Thr Ser Pro
 450 455 460
 Lys Arg Trp Ile Tyr Asp Thr Ser Lys Leu Ala Ser Gly Val Pro Ala
 465 470 475 480
 His Phe Arg Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser
 485 490 495
 Gly Met Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser
 500 505 510
 Ser Asn Pro Phe Thr Phe Gly Ser Gly Thr Lys Leu Glu Ile Asn Arg
 515 520 525
 Ala Asp Thr Ala Pro Thr Gly Ser Glu Gln Lys Leu Ile Ser Glu Glu
 530 535 540
 Asp Leu Asn Ser His His His His His
 545 550

(2) INDICATIONS AS TO ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1653 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: genome DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) POSITION: 28..1644

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) POSITION: 28..1644

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAATTCAATTA AAGAGGAGAA ATTAACC ATG AAA TAC CTA TTG CCT ACG GCA Met Lys Tyr Leu Leu Pro Thr Ala	51
1 5	
GCC GCT GGC TTG CTG CTG GCA GCT CAG CCG GCC ATG GCG CAG GTG Ala Ala Gly Leu Leu Leu Ala Ala Gln Pro Ala Met Ala Gln Val	99
10 15 20	
CAA CTG CAG CAG TCT GGG GCT GAA CTG GCA AGA CCT GGG GCC TCA GTG Gln Leu Gln Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val	147
25 30 35 40	
AAG ATG TCC TGC AAG GCT TCT GGC TAC ACC TTT ACT AGG TAC ACG ATG Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Arg Tyr Thr Met	195
45 50 55	
CAC TGG GTA AAA CAG AGG CCT GGA CAG GGT CTG GAA TGG ATT GGA TAC His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile Gly Tyr	243
60 65 70	
ATT AAT CCT AGC CGT GGT TAT ACT AAT TAC AAT CAG AAG TTC AAG GAC Ile Asn Pro Ser Arg Gly Tyr Thr Asn Tyr Asn Gln Lys Phe Lys Asp	291
75 80 85	
AAG GCC ACA TTG ACT ACA GAC AAA TCC TCC AGC ACA GCC TAC ATG CAA Lys Ala Thr Leu Thr Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln	339
90 95 100	
CTG AGC AGC CTG ACA TCT GAG GAC TCT GCA GTC TAT TAC TGT GCA AGA Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg	387
105 110 115 120	
TAT TAT GAT GAT CAT TAC AGC CTT GAC TAC TGG GGC CAA GGC ACC ACT Tyr Tyr Asp His Tyr Ser Leu Asp Tyr Trp Gly Gln Gly Thr Thr	435
125 130 135	
CTC ACA GTC TCC TCA GCC AAA ACA ACA CCC AAG CTT GGC GGT GAT ATC Leu Thr Val Ser Ser Ala Lys Thr Thr Pro Lys Leu Gly Gly Asp Ile	483
140 145 150	
TTG CTC ACC CAA ACT CCA GCT TCT TTG GCT GTG TCT CTA GGG CAG AGG Leu Leu Thr Gln Thr Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg	531
155 160 165	
GCC ACC ATC TCC TGC AAG GCC AGC CAA AGT GTT GAT TAT GAT GGT GAT Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val Asp Tyr Asp Gly Asp	579
170 175 180	

AGT TAT TTG AAC TGG TAC CAA CAG ATT CCA GGA CAG CCA CCC AAA CTC Ser Tyr Leu Asn Trp Tyr Gln Gln Ile Pro Gly Gln Pro Pro Lys Leu 185 190 195 200	627
CTC ATC TAT GAT GCA TCC AAT CTA GTT TCT GGG ATC CCA CCC AGG TTT Leu Ile Tyr Asp Ala Ser Asn Leu Val Ser Gly Ile Pro Pro Arg Phe 205 210 215	675
AGT GGC AGT GGG TCT GGG ACA GAC TTC ACC CTC AAC ATC CAT CCT GTG Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Asn Ile His Pro Val 220 225 230	723
GAG AAG GTG GAT GCT GCA ACC TAT CAC TGT CAG CAA AGT ACT GAG GAT Glu Lys Val Asp Ala Ala Thr Tyr His Cys Gln Gln Ser Thr Glu Asp 235 240 245	771
CCG TGG ACG TTC GGT GGA GGC ACC AAG CTG GAA ATC AAA CGG GCT GAT Pro Trp Thr Phe Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Asp 250 255 260	819
GCT GCG GCC GCT GGT GGC CCA GGG TCG CAG GTG CAG CTG CAG CAG TCT Ala Ala Ala Ala Gly Gly Pro Gly Ser Gln Val Gln Leu Gln Gln Ser 265 270 275 280	867
GGG GCT GAG CTG GTG AGG CCT GGG TCC TCA GTG AAG ATT TCC TGC AAG Gly Ala Glu Leu Val Arg Pro Gly Ser Ser Val Lys Ile Ser Cys Lys 285 290 295	915
GCT TCT GGC TAT GCA TTC AGT AGC TAC TGG ATG AAC TGG GTG AAG CAG Ala Ser Gly Tyr Ala Phe Ser Ser Tyr Trp Met Asn Trp Val Lys Gln 300 305 310	963
AGG CCT GGA CAG GGT CTT GAG TGG ATT GGA CAG ATT TGG CCT GGA GAT Arg Pro Gly Gln Gly Leu Glu Trp Ile Gly Gln Ile Trp Pro Gly Asp 315 320 325	1011
GGT GAT ACT AAC TAC AAT GGA AAG TTC AAG GGT AAA GCC ACT CTG ACT Gly Asp Thr Asn Tyr Asn Gly Lys Phe Lys Gly Lys Ala Thr Leu Thr 330 335 340	1059
GCA GAC GAA TCC TCC AGC ACA GCC TAC ATG CAA CTC AGC AGC CTA GCA Ala Asp Glu Ser Ser Thr Ala Tyr Met Gln Leu Ser Ser Leu Ala 345 350 355 360	1107
TCT GAG GAC TCT GCG GTC TAT TTC TGT GCA AGA CGG GAG ACT ACG ACG Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg Arg Glu Thr Thr 365 370 375	1155
GTA GGC CGT TAT TAC TAT GCT ATG GAC TAC TGG GGT CAA GGA ACC TCA Val Gly Arg Tyr Tyr Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr Ser 380 385 390	1203
GTC ACC GTC TCC TCA GCC AAA ACA ACA CCC AAG CTT GGC GGT GAT ATC Val Thr Val Ser Ser Ala Lys Thr Thr Pro Lys Leu Gly Gly Asp Ile 395 400 405	1251

GTG CTC ACT CAG TCT CCA GCA ATC ATG TCT GCA TCT CCA GGG GAG AAG Val Leu Thr Gln Ser Pro Ala Ile Met Ser Ala Ser Pro Gly Glu Lys 410 415 420	1299
GTC ACC ATG ACC TGC AGT GCC AGC TCA AGT GTA AGT TAC ATG AAC TGG Val Thr Met Thr Cys Ser Ala Ser Ser Val Ser Tyr Met Asn Trp 425 430 435 440	1347
TAC CAG CAG AAG TCA GGC ACC TCC CCC AAA AGA TGG ATT TAT GAC ACA Tyr Gln Gln Lys Ser Gly Thr Ser Pro Lys Arg Trp Ile Tyr Asp Thr 445 450 455	1395
TCC AAA CTG GCT TCT GGA GTC CCT GCT CAC TTC AGG GGC AGT GGG TCT Ser Lys Leu Ala Ser Gly Val Pro Ala His Phe Arg Gly Ser Gly Ser 460 465 470	1443
GGG ACC TCT TAC TCT CTC ACA ATC AGC GGC ATG GAG GCT GAA GAT GCT Gly Thr Ser Tyr Ser Leu Thr Ile Ser Gly Met Glu Ala Glu Asp Ala 475 480 485	1491
GCC ACT TAT TAC TGC CAG CAG TGG AGT AGT AAC CCA TTC ACG TTC GGC Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Ser Asn Pro Phe Thr Phe Gly 490 495 500	1539
TCG GGG ACA AAG TTG GAA ATA AAC CGG GCT GAT ACT GCA CCA ACT GGA Ser Gly Thr Lys Leu Glu Ile Asn Arg Ala Asp Thr Ala Pro Thr Gly 505 510 515 520	1587
TCC GAA CAA AAG CTG ATC TCA GAA GAA GAC CTA AAC TCA CAT CAC CAT Ser Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ser His His His 525 530 535	1635
CAC CAT CAC TAATCTAGA His His His	1653

(2) INDICATIONS AS TO ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 539 amino acids
 - (B) KIND: amino acid
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala 1 5 10 15
Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Gln Ser Gly Ala Glu 20 25 30
Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly 35 40 45

Tyr Thr Phe Thr Arg Tyr Thr Met His Trp Val Lys Gln Arg Pro Gly
 50 55 60

Gln Gly Leu Glu Trp Ile Gly Tyr Ile Asn Pro Ser Arg Gly Tyr Thr
 65 70 75 80

Asn Tyr Asn Gln Lys Phe Lys Asp Lys Ala Thr Leu Thr Thr Asp Lys
 85 90 95

Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp
 100 105 110

Ser Ala Val Tyr Tyr Cys Ala Arg Tyr Tyr Asp Asp His Tyr Ser Leu
 115 120 125

Asp Tyr Trp Gly Gln Gly Thr Thr Leu Thr Val Ser Ser Ala Lys Thr
 130 135 140

Thr Pro Lys Leu Gly Gly Asp Ile Leu Leu Thr Gln Thr Pro Ala Ser
 145 150 155 160

Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser
 165 170 175

Gln Ser Val Asp Tyr Asp Gly Asp Ser Tyr Leu Asn Trp Tyr Gln Gln
 180 185 190

Ile Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr Asp Ala Ser Asn Leu
 195 200 205

Val Ser Gly Ile Pro Pro Arg Phe Ser Gly Ser Gly Thr Asp
 210 215 220

Phe Thr Leu Asn Ile His Pro Val Glu Lys Val Asp Ala Ala Thr Tyr
 225 230 235 240

His Cys Gln Gln Ser Thr Glu Asp Pro Trp Thr Phe Gly Gly Thr
 245 250 255

Lys Leu Glu Ile Lys Arg Ala Asp Ala Ala Ala Gly Gly Pro Gly
 260 265 270

Ser Gln Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly
 275 280 285

Ser Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ala Phe Ser Ser
 290 295 300

Tyr Trp Met Asn Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp
 305 310 315 320

Ile Gly Gln Ile Trp Pro Gly Asp Gly Asp Thr Asn Tyr Asn Gly Lys
 325 330 335

Phe Lys Gly Lys Ala Thr Leu Thr Ala Asp Glu Ser Ser Ser Thr Ala
 340 345 350
 Tyr Met Gln Leu Ser Ser Leu Ala Ser Glu Asp Ser Ala Val Tyr Phe
 355 360 365
 Cys Ala Arg Arg Glu Thr Thr Val Gly Arg Tyr Tyr Ala Met
 370 375 380
 Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Lys Thr
 385 390 395 400
 Thr Pro Lys Leu Gly Gly Asp Ile Val Leu Thr Gln Ser Pro Ala Ile
 405 410 415
 Met Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Ser Ala Ser
 420 425 430
 Ser Ser Val Ser Tyr Met Asn Trp Tyr Gln Gln Lys Ser Gly Thr Ser
 435 440 445
 Pro Lys Arg Trp Ile Tyr Asp Thr Ser Lys Leu Ala Ser Gly Val Pro
 450 455 460
 Ala His Phe Arg Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
 465 470 475 480
 Ser Gly Met Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
 485 490 495
 Ser Ser Asn Pro Phe Thr Phe Gly Ser Gly Thr Lys Leu Glu Ile Asn
 500 505 510
 Arg Ala Asp Thr Ala Pro Thr Gly Ser Glu Gln Lys Leu Ile Ser Glu
 515 520 525
 Glu Asp Leu Asn Ser His His His His His His
 530 535

(2) INDICATIONS AS TO ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 57 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TATATACTGC AGCTGCACCT GCGACCCTGG GCCACCAGCG GCCGCAGCAT CAGCCCG

57

(2) INDICATIONS AS TO ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear

- (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CCGTGAATT CAGGTGCAAC TGCAGCAGTC TGGGGCTGAA CTGGC

45

(2) INDICATIONS AS TO ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear

- (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGTCGACGTT AACCGACAAA CAACAGATAA AACG

34

(2) INDICATIONS AS TO ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 348 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: genome DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) POSITION: 1..348
- (ix) FEATURE:
 - (A) NAME/KEY: mat_peptide
 - (B) POSITION: 1..348
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA TCC TCC	48
Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser	
1 5 10 15	
GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA CAA	96
Ala Leu Ala Ala Pro Val Asn Thr Thr Glu Asp Glu Thr Ala Gln	
20 25 30	
ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC	144
Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe	
35 40 45	
GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG	192
Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu	
50 55 60	
TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA	240
Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val	
65 70 75 80	
TCT CTC GAG AAA AGA GAG GCT GAA GCT GAA TTC CAG GTG CAA CTG CAG	288
Ser Leu Glu Lys Arg Glu Ala Ala Glu Phe Gln Val Gln Leu Gln	
85 90 95	
CAG TCT GGG GCT GAA CTG GCA AGA CCT GGG GCC TCA GTG AAG ATG TCC	336
Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser	
100 105 110	
TGC AAG GCT TCT	348
Cys Lys Ala Ser	
115	

2) INDICATIONS AS TO ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 116 amino acids
 - (B) KIND: amino acid
 - (D) TOPOLOGY: linear

(ii) KIND OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1 5 10 15

Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
..... 20 25 30

Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
 35 40 45

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
50 55 60

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
65 70 75 80

Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Gln Val Gln Leu Gln
85 90 95

Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser
100 105 110

Cys Lys Ala Ser
115

(2) INDICATIONS AS TO ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 354 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear

(ii) KIND OF MOLECULE: genome DNA

(iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

(ix) FEATURES

- (A) NAME/KEY: CDS
(B) POSITION: 1 354

(B) POS

- (A) NAME/KEY: mat_peptide
(B) POSITION: 1-254

(B) POSITION: 1..354

ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA TCC TCC Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser 1 5 10 15	48
GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA CAA Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln 20 25 30	96
ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe 35 40 45	144
GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu 50 55 60	192
TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val 65 70 75 80	240
TCT CTC GAG AAA AGA GAG GCT GAA GCT GAA TTC ATG GCG CAG GTG CAA Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Met Ala Gln Val Gln 85 90 95	288
CTG CAG CAG TCT GGG GCT GAA CTG GCA AGA CCT GGG GCC TCA GTG AAG Leu Gln Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys 100 105 110	336
ATG TCC TGC AAG GCT TCT Met Ser Cys Lys Ala Ser 115	354

2) INDICATIONS AS TO ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 118 amino acids
 - (B) KIND: amino acid
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser 1 5 10 15
Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln 20 25 30
Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe 35 40 45

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
 50 55 60

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 65 70 75 80

Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Met Ala Gln Val Gln
 85 90 95

Leu Gln Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys
 100 105 110

Met Ser Cys Lys Ala Ser
 115

(2) INDICATIONS AS TO ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 42 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TCACACAGAA TTCTTAGATC TATTAAAGAG GAGAAATTAA CC

42

(2) INDICATIONS AS TO ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

AGCACACGAT ATCACCGCCA AGCTTGGGTG TTGTTTGCG

40

(2) INDICATIONS AS TO ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 43 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14

AGCACACAAG CTTGGCGGTG ATATCTTGCT CACCCAAACT CCA

43

(2) INDICATIONS AS TO ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 57 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15

AGCACACTCT AGAGACACAC AGATCTTAG TGATGGTGAT GGTGATGTGA GTTTAGG

57

(2) INDICATIONS AS TO ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear

- (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CAGCCGGCCA TGGCGCAGGT GCAACTGCAG CAG

33

(2) INDICATIONS AS TO ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 102 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

TATATACTGC AGCTGCACCT GGCTACCACC ACCACCGGAG CCGCCACCAC CGCTACCACC

60

GCCGCCAGAA CCACCAACAC CAGCGGCCGC AGCATCAGCC CG

102